MONOGRAPH 6

RECOGNITION AND IDENTIFICATION OF OBJECTS

--- A Theory of Inference ----

Lee Roy Beach

SAM-P. 18

U. S. NAVAL SCHOOL OF AVIATION MEDICINE
U. S. NAVAL AVIATION MEDICAL CENTER
PENSACOLA FLORIDA

ACKNOWLEDGEMENTS

The author would like to express sincere appreciation to Prof. K. R. Hammond, Prof. M. Wertheimer, Dr. M. B. Jones, Mr. J. R. Berkshire, and especially to Dr. D. E. Bailey for their valuable advice and encouragement. Particular thanks are due Mrs. Alice Garman for her patience and stamina and to the U.S. Naval School of Aviation Medicine for its support. The opinions expressed herein are not to be construed as necessarily reflecting the views or endorsement of the Navy Department.

U.S. NAVAL SCHOOL OF AVIATION MEDICINE

Monograph Series, No. 6

A THEORY OF INFERENCE

by
Lee Roy Beach, Ph.D.
LT, MSC, USNR

Approved by:	Released by:
Captain Ashton Graybiel	Captain Clifford P. Phoebus
MC, USN	MC, USN
Director of Research	Commanding Officer

18 March 1963

TABLE OF CONTENTS

	Page
PREFACE	ii
THE CONCEPT OF INFERENCE	1
ATTRIBUTES OF OBJECTS	3
Classes	4
Cue Dimensions	9
THE COGNITIVE SPACE	11
INFERENCE AND THE COGNITIVE SPACE	13
Inference of Class Membership	13
Inference of Cue Values	25
INFERENTIAL ACCURACY	32
The Revision Process	32
GENERAL REMARKS	39
REFERENCES	43
NOTES	48

PREFACE

This theory got its start in 1959 when I became engrossed in Egon Brunswik's Theory of Probabilistic Functionalism (1956). As my familiarity with his theory increased, however, it became apparent that while it provided a profitable and unique way of viewing behavior it was not a formal theory. Rather, because of his premature death, Brunswik's approach was an incomplete framework within which further theorizing and research were needed before an adequate theory could evolve.

That this feeling was shared by others was evidenced by Bruner, Goodnow, and Austin's _A Study of Thinking_ (1956) and Sarbin, Taft, and Bailey's _Clinical Inference and Cognitive Theory_ (1960). Both of these books extend and develop the Brunswikian line of thought while underscoring the essential continuity underlying perception, cognition, and inference. The present theory is in the same vein.

At the time the theory was beginning to approach a stable form the book by Sarbin, _et al._, was published. Their Theory of Modular Organization and my own thinking were highly similar, which is not surprising in view of their common origins in Brunswik's work. However, the theories were not identical and the differences between them increased when a new tack was selected for development of the present theory. In order to extend the range of theoretical coverage an attempt was made to treat in detail some of the aspects of inference which were only touched upon by Sarbin, _et al._ Thus the emphasis in the present work is upon the role of inference in the acquisition of knowledge about objects. In the other theory the emphasis is upon the use of inference in planning interactions with the object, particularly in the clinical situation. It would appear, however, that the origin of the knowledge used in subsequent interaction has a claim to prior theoretical importance.

In the present work an attempt has been made to

utilize fully the mathematical properties of the cognitive space. In doing so it was necessary to adopt a more stringent definition and analysis of classes than was presented in the Modular Organization Theory. This emphasis on the metric properties of the cognitive space also permits the present theory more precise predictions about subjects' inference behavior and has allowed rather detailed elaboration of various theoretical mechanisms and their hypothetical relationships to other of the constructs.

Perhaps the greatest difference between the two theories is in how they handle the individuality of the object. If I understand correctly, in the Modular Organization Theory the object's known attributes are used to relegate it to a class, i.e., are used to determine what kind of object it is. Then, however, these unique attributes are ignored and the inference about the object is based upon the properties of the class as a whole. For example, a particular object exhibits certain symptoms and is classified as a schizophrenic. Then his behavior is predicted entirely upon the knowledge the subject possesses about the class of objects called schizophrenics. I doubt, however, if this is actually what people do. In the present theory the object's class is important in determining the inferences about it but its own <u>individual properties</u> also influence the final decision.

This difference between the two theories has not yet been successfully tested but it is indicative of the fact that testable differences do exist. The theories are not wholly alike but, rather, they are complimentary. Where they overlap some differences exist, differences which can be empirically decided and which will thereby contribute to a stronger, broader theory of inference and knowledge.

Pensacola, Florida
March 1963

Lee Roy Beach

THE CONCEPT OF INFERENCE

The course of a man's life is intricately interwoven with the objects--persons, things, and events--which he encounters and with which he interacts. Over a lifetime millions of different objects are experienced, some repeatedly and some only once. In view of their number and variety, the feat of successful identification or recognition of these objects and the appropriate interaction with them deserves the close attention of students of human behavior.

Interaction with objects (*Os*) presumes knowledge about them, knowledge which allows the subject (*S*) to plan his part of the interaction properly and to have some notion of how the *O* will react to his actions. It is only through the proper use of knowledge about *Os* in the planning of interactions that *S* can obtain the social, physical, and intellectual goals which sustain life and which make it agreeable and interesting.

Knowledge about *Os* can be divided into two general categories. The first is knowledge about the kind of *O* it is, i.e., the general class of *Os* to which it belongs--a house, a dog, a book, a man, or the like. The second is knowledge about the nature of the *O*, i.e., the *O*'s properties or characteristics--its intelligence, color, size, friendliness, texture, and so on. When an *O* is first encountered, *S* obtains some information about it through immediately apparent characteristics, such as contour, color, and sound. From this scanty information he attempts to determine the kind of *O* it is. Then the knowledge about what it is, together with the knowledge about its immediately apparent characteristics, can be used to make educated guesses about the nature of the *O*. The present work is interested in how *Ss* utilize information about the know attributes of *Os* in order to make inferences about the *Os*' unknown attributes.

Inference, as a psychological concept, has too frequently been limited solely to the realm of cognitive functioning. As we shall see, however, the wider application of this concept, with concomitant specification of the mechanisms involved, can prove valuable. The continuity from the chaos of multiple sensory events to orderly knowledge about *O*s can be viewed as a series of inferences. Since there is no single sensory event signifying an *O*'s intelligence, for example, it must be an inference based upon multiple sensory events and upon other, previously inferred, knowledge about the *O*. Inference can account for the fact than an *O* can be recognized repeatedly, even though the *O* seldom presents exactly the same patterns of sensory stimulation. The concept of inference can also account for *S*'s ability to determine the identity of an *O* that he has never seen before. It can also account for how *S*s can make fairly shrewd judgments about the nature of an *O* solely on the basis of knowing what kind of *O* it is. In short, the broader application of this concept provides a common strand through a number of different psychological events and thereby may aid in placing these events within a unifying theoretical framework.

This article is devoted to the construction of a theory of inference. The general paradigm for the theory is that *S* encounters *O*s which possess various attributes. These attributes are values on various cue dimensions and class memberships. On each encounter some of the *O*'s cue values are given as immediately apparent aspects of the *O*. From these *S* must make inferences about what kind of *O* he has encountered (i.e., the *O*'s class) and, subsequently, about those cue values for which no information has yet been received.

In brief, the theory assumes that each *S* possesses a multidimensional cognitive space which is defined by the cue dimensions to which he attends. All of the cue values *S* knows to be possessed by the *O* define a location in this space which comes to be the cognitive representation of the *O*, called *θ*,

which lasts over time. If, when an *O* is referred to the space, a *Θ* already exists at or near the location, the *O* is said to be "*recognized*" as the previously experienced *O* corresponding to the *Θ*. Then the cue values and the class associated with the *Θ* can be assumed to be applicable to the recognized *O*. If the location defined by the encountered *O*'s cue values falls near but not at a *Θ* in the space, *S* can still assume the *Θ* and *O* to be highly similar in class membership and in covert cue values. Here the class and covert cue values associated with the similar *Θ* can be "*assimilated*" for the newly encountered *O* and assumed to be sufficient until further information indicates the contrary. If the location defined by the encountered *O*'s cue values does not fall close to a *Θ* in the space, *S* can fall back on a third method of inferring the *O*'s class and cue values. This third method, called "*identification*", relies on the statistical properties of groups of *Θ*s which possess cue values similar to the ones the *O* is known to possess. This process, together with the other two, comprises the major portion of the following discussion. The remainder of the paper concerns the methods by which the inferences are checked against subsequent information and how revisions are made in an effort to maintain cognitive accuracy.

ATTRIBUTES OF OBJECTS

In their *Study of Thinking*, Bruner, Goodnow and Austin (1956) define an attribute of an *O* as "any discriminable feature of an *O* that is susceptible of some discriminable variation from *O* to *O*." These authors also quote Boring's (1942) similar definition: "A stone is shape, color, weight, and kind of substance in complicated relation. When such descriptive ultimates are general properties which can vary continuously or discretely, when they are, in short, parameters, they may, if one chooses, be called attributes of the object described."

As these definitions indicate, *O*s are defined by their attributes and attributes are those aspects of

Os which permit discriminations among them. It is assumed that there are two types of attributes accociated with Os, classes and cue values.

Classes

The formal definition of a class is that it is an attribute of Os which permits only dichotomous (presence or absence) judgments about its association with any one O. As such, a class is a nominal attribute of an O which does not permit discriminations among the Os which possess it. For example, knowing that two Os are members of the class "cow" does not aid in discriminating between them--discriminations are based upon other attributes such as color, size, and markings. The discriminating attributes are cue values.

On a less formal level the definiation of an O's class can be stated as what kind of O it is as opposed to what its properties are. An O is a man (class) who is of exceptional height (cue value). Classes are impalpable attributes of Os--attributes which exist as a function of the psychological operation of putting Os together into various groups and of assigning these groups specific names. This does not mean that classes cannot be subdivided on the basis of discriminations made on cue dimensions, e.g., red houses and white houses, it merely means that the class remains the group of Os to which the discrimination refers. At no time, however, can the members of one class be discriminated from one another on the basis of knowledge of the class alone.

As we shall see in the course of the subsequent discussion, the theory assumes S uses an O's immediately given cue value to infer its class and then goes from the class to further inferences about the O's unknown cue values. It is clear that it is logically possible to construct a theory in which the inference goes directly from the O's known cue values to the unknown cue values. Psychologically, however, the validity of the second approach can be questioned.

Inferences are always about things--which implies that the O is part of a limited group which has a name. It is only when S knows what the O is that it becomes clear which covert cue values are relevant to the ensuing interaction--certainly the known height of an O means different things depending on what the O is. (To cite an extreme example, a man who is five feet tall will prompt a set of inferences about his properties and a course of interaction which are quite different from those prompted by a cat of equal height.) The psychologically meaningful function of classes as intermediaries in the inference of unknown cue values can be taken as evidence for the necessity of differentiating between classes and cues in the present theory.

The Class Hierarchy. It is obvious that an O is associated with many different classes. Moreover, it is equally true that knowledge of its membership in one class reveals its membership in a number of other classes. This is because classes exist in a hierarchy. The hierarchy runs from an infinite number of small classes (each containing only one O) at the bottom on up through classes which are fewer and fewer in number but increasingly larger in size. At each level of the class hierarchy the classes are all mutually exclusive.

Any particular O is the member of a "chain" of classes extending from the lowest levels of the hierarchy to the highest. At each level the class, or part of the class, combines with other classes, or parts of classes, on the same level to form another, higher level, class. For example, an O may be a "cottage" which can be combined with the classes, "apartment", "house", "mansion", etc., and subsumed under the higher-order class "residence". This in turn can be combined with "store", "depot", "bank", and the like to form the higher-order class "building". "Building", "stadium", "fort", etc., form the yet higher-order class "structure", and on and on.

Inference and the Class Hierarchy Level. When

one of the O's class memberships is known, its memberships at many of the higher levels in the hierarchy are also known; however, the O's class memberships for lower levels are still unknown. Thus, if S knows an O is a residence he also knows it is a building and a structure. But, he does not know if the residence is a cottage, an apartment house, or a mansion.

Consequently, when an unfamiliar O is encountered S must select a level in the class hierarchy at which to make the inference. While the lowest level possible would be best, because this would usually reveal its membership in a number of higher-order classes, this is seldom possible. To classify an O as a member of a very low level class it is necessary to know a good deal about the O's attributes and S seldom possesses much information at the moment he is forced to make his decision. Given a limited amount of information, the selection of a higher level at which to work will usually increase the likelihood of a correct inference. For example, there is a smaller possibility of error in assigning an O to the broad class "structure" than in deciding it is a member of the smaller, more tightly defined class, "cottage".

On the other hand, after S decides what the O is he usually intends to make further inferences about its cue values. These cue inferences are based on knowledge S has about the O's fellow class members, and, as a result, he does best to select a lower-order class initially because such classes are smaller and are composed of more homogeneous Os. In general then, the higher the order of the class the more likely is the identification of the O to be correct; the lower the order of the class the more likely are subsequent cue inferences to be correct. The class level at which these two factors balance one another will be the one selected for the identification of an unfamiliar O.

Classes and the Cognitive Space. In order to fit classes into the framework of the cognitive space it is necessary to make a special assumption about them. This assumption is that at each level of the class hierarchy the elements which make up the cognitive space are partitioned into classes. This is accomplished by a class function which is merely a counting measure on the space. The hypothetical structure of the partitioned cognitive space for each level in the class hierarchy is illustrated in Fig. 1. As this diagram shows, higher-order classes are composites of wholes and parts of lower-order classes. (There is no empty class at any level because every O must belong to one of the classes.)

The set of classes which comprise the class function are the "presence" ends of the previously discussed dichotomous class dimensions. Any existing order among the classes is imposed by the higher-order classes, the members of subsumed classes are ordinarily somewhat similar because the origin of classes in social agreement and linguistic naming practices usually is predicated upon the similarity among Os. (See Bruner, Goodnow and Austin, 1956, and Miller and Dollard, 1941, for a discussion of the origin of classes.) The necessity for the assumption of this class function will become apparent as the discussion progresses, particularly when the structure of the cognitive space is examined.

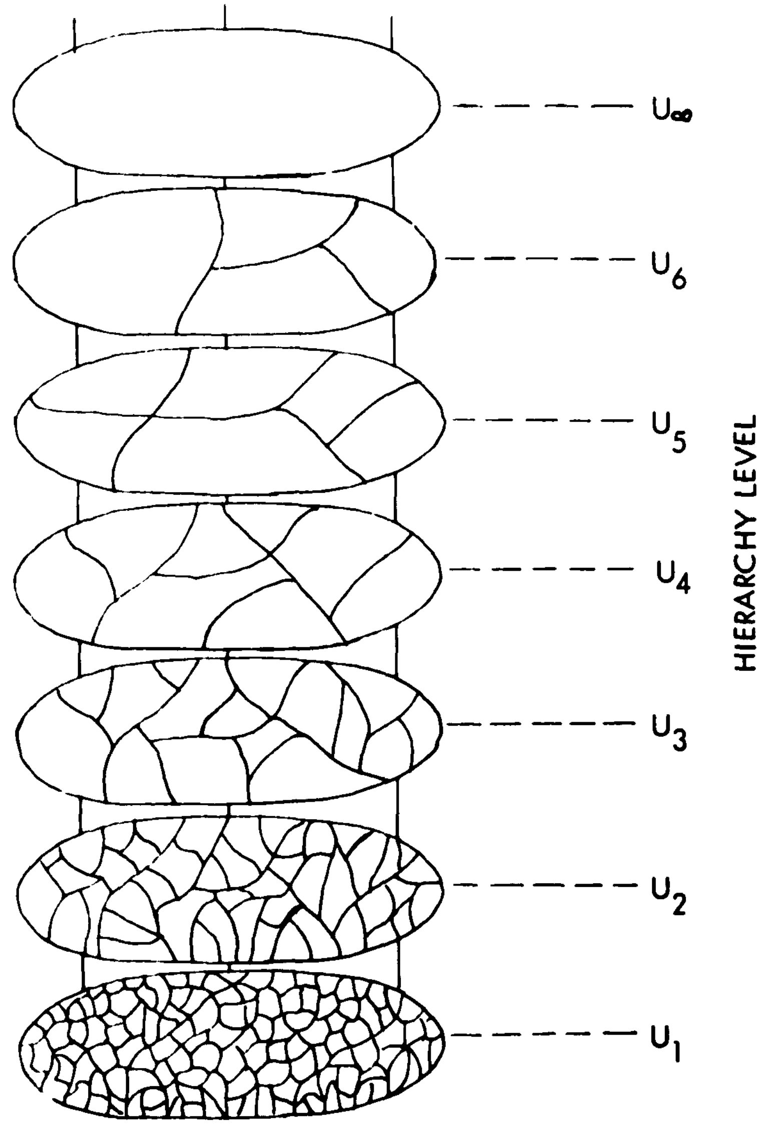

Fig. 1. The hypothetical structure of the cognitive space at various levels of the class hierarchy. At the lowest level each class contains one $\underline{\Theta}$ and at the highest level one class contains all $\underline{\Theta}$s.

Figure 1

Cue Dimensions

The classes are formally defined as attributes for which only two values exist, presence or absence. For some attributes, however, it is possible not only to discriminate among Os on the basis of merely the presence or absence of an attribute, but it is also possible to discriminate among them on the basis of the degree to which the attribute is present. For example, some Os not only possess the attribute of weight but some are heavier or lighter than others. Similarly, some Os are happier than others, or harder, or less friendly, or wider, or smaller, or of rougher texture, etc. When such discriminations can be made among Os possessing an attribute it qualifies as a cue dimension and the degree to which each O possesses one of these attributes is the O's value on the cue dimension.

Cue dimensions, like other continuous measuring scales, potentially possess an infinite number of values, that is, each cue value can be subdivided into a number of smaller segments of the dimension, and each of these can in turn be subdivided, etc. It is doubtful, however, whether Ss actually fully utilize this property of the dimensions. In the first place the S's receptor threshold places lower bounds on his ability to divide a dimension into small segments. In the second place Ss tend to use a coarser metric than they are actually capable of discriminating; when you meet a man on the street he is not 6'3", he is "tall". It would appear that, given the limits imposed by the receptor threshold and the situation (and the latter also influences the former, Swets, 1961), Ss are able to consolidate and differentiate the values on the cue dimensions thereby suiting the coarseness of the metric to demands of the situation.

Psychologically, cues are differentiated from classes by the fact that, on one hand, they are the attributes of Os which are immediately apparent when an O is encountered; the bases for the inference

process. On the other hand, those attributes of the O which are unknown when the O is first encountered and which are directly relevant to the success of the ensuing interation are also cues. That is, knowing an O is a chair (its class) is not directly relevant to interaction with the O, but knowing that chairs usually have the properties of being solid, are 1 1/2 feet from floor to seat, that the seat is parallel with the floor, etc., are all cue values which guide S's interaction with the O. In other words the O is a chair and has various attributes--attributes which, by the way, are not necessarily the same for all members of the class. (A rule of thumb for distinguishing cues and classes is to construct an English sentence about them. Proper nouns, common nouns, collective nouns, and personal pronouns are classes of Os while abstract nouns, adjectives, participles, and adverbs are usually cue values associated with the Os.)

The Cue Dimension Hierarchy. Just as classes exist in a hierarchy so is there a hierarchy of cue dimensions. The cue dimension hierarchy extends from the higher-order, culturally determined attributes, like friendliness, on down to the sensory level. (At all levels, however, the cue dimensions are multivalued.) Clearly, few of the attributes we associate with Os are sensorily detected. For most meaningful cue dimensions the Os' values are only remotely related to the original sensory stimulation. From the present point of view, stimuli (sound, light, temperature, etc.) emanating from the O stimulate S's receptors and produce sensations of various degrees of loudness, dark, light, color, heat, etc. These lowest-order cues are then used to infer slightly higher-order cues such as edge, texture, surface, size, and the like. Next, these cues lead to the inference of even higher-order cues--square corners, planes of surface, orientation in space, extension, and so on up the cue dimension hierarchy. In a sense, S "generates" information about the O in terms of inferences about increasingly higher-order cue values.

In the course of inference from lower-order to higher-order cue values S arrives at a point at which he has enough information to permit inference of the O's class. While the accuracy of the class inference is likely to increase with increased information about the O, the S is usually at great pains to infer the O's class as soon as possible. This is because until the O's class is known S cannot proceed to make higher-order inferences. Cue values such as friendliness, distance, size, etc., are dependent not only upon sensory stimulation from the O but also upon knowledge of what the O is. (See, for example, Ames, 1955; Bruner and Minturn, 1955; Bruner, Postman, and Rodrigues, 1951; Ittleson and Slack, 1958.) Therefore, S must make his class inference before proceeding to infer higher-order cue values. The class inference will occur when S possesses the minimum safe amount of lower-order cue information about the O, i.e., when he is relatively confident that his inference will be correct. Then he is able to utilize the class in subsequent inferences about the O's higher-order cue values.

The Class and the Cue Dimension Hierarchy. The similarity between the class hierarchy and the cue dimension hierarchy is not so great as it at first appears. The class hierarchy, as illustrated in Fig. 1, is composed of sets of Os which are repartitioned at every level of the hierarchy. As a result, knowledge of an O's lower-order class often reveals its membership in a number of the classes on the higher levels. The cue hierarchy does not consist of a redefined set of the same elements at each level. The levels of this hierarchy only indicate the abstractness or "remoteness" (Brunswik, 1956) of the various cue dimensions. The O's value on each succeeding higher-order cue dimension must be inferred from the lower-order cue values.

THE COGNITIVE SPACE

The assumed structure of the cognitive space is similar to that of analagous concepts proposed by Kelly (1955); Osgood, Suci, and Tannenbaum (1957);

Overall and Williams (1961); Sarbin, Taft, and Bailey (1960), and others. It is a multidimensional space which is defined by the cue dimensions. Cells are formed by the intersection of the cue values with other cue values.

The cognitive space provides an information storage and "retrieval" system for the Os which S encounters. Each cue value associated with an O can be located on a cue dimension bounding S's cognitive space. All of these values define a particular cell in the space which is the key both to retention of information about the O and to retrieval of this knowledge when it is needed. The cell serves to retain knowledge because its location is a succinct summary of everything S knows about an O. It serves in the retrieval of this knowledge because, through its location, specification of the cell reveals all of the cue values which S knows for the O.

Because it represents each of the O's cue values, the cell is unique to that O and the cell is its "cognitive representation" (if two Os have exactly the same attributes they cannot be discriminated and are functionally the same O). This cell is designated by the symbol Θ (Theta).

The class hierarchy fits into the concept of the cognitive space as a set of measure functions which partition the universe of Θs into sets or classes. The number of classes is, of course, determined by the level (U in Fig. 1) of the hierarchy at which S has decided to operate (cf. p. 6). This means that, in addition to the cue values which locate it in the cognitive space, the Θ also has associated with it the class to which it belongs. And, just as its Θ serves as a retrieval system for the O's cue values, so too can the Θ be used to gain access to the O's class membership.

The structure of the cognitive space makes it amenable to many of the mathematical operations which are familiar to psychologists. For example, it is

possible to sum over dimensions to obtain distributions of $\underline{\Theta}$s for other, specified dimensions. Or, the $\underline{\Theta}$s in a particular area of the space can be singled out for scrutiny, as for example a single class of $\underline{\Theta}$s, or only those $\underline{\Theta}$s which belong to a given class and also possess given values on specified cue dimensions. These properties of the space underlie S's utilization of the $\underline{\Theta}$s which lie within it as the basis for inferences.

INFERENCE AND THE COGNITIVE SPACE

S bases his interaction with an O upon knowledge about its class and its cue values. As a result, a good deal of effort is devoted to acquiring this knowledge and to storing it away in the cognitive space for future use. We now turn to the various ways of using this knowledge for making inferences upon which the interaction can be based.

INFERENCE OF CLASS MEMBERSHIP

Recognition. When an O appears before S, only a few lower-level cues are apparent. Using these cues, S's first task is to decide whether or not this O, whatever it may be, has ever been experienced before; i.e., if it is familiar or unfamiliar. If it is recognized as familiar, a great deal may already be known about it and this knowledge can be utilized in the ensuing interaction. Because of changes in situational conditions, however, as well as changes in an O over time, the basic cue values associated with an O may be slightly different from one encounter to the next. S's immediate problem is to decide whether the difference in cue values exhibited by the O being encountered at the moment and those possessed by an O seen previously is large enough to indicate that they are not the same O. Or, stated in terms of the cognitive space, recognition boils down to deciding whether the cue values associated with a $\underline{\Theta}$ in the space are sufficiently similar to those associated with the present O to permit the assumption that the $\underline{\Theta}$ derives from a previous encounter with this same O.

To make the recognition inference the O's immediately apparent cue values are used to define a cell in the cognitive space, called Ø. Next a measure is made of the distance between the Ø and every θ in the cognitive space.[1] The smaller the distance between the Ø and the nearest θ, the more likely S is to assume that they both derive from the same O. There is, however, some critical distance beyond which S will reject the hypothesis that both the Ø and the nearest θ derive from the same O and therefore the O is not recognized.[2] (See Ames, 1949, and Arnoult, 1956, for research concerning Ss' assumption of identity between sufficiently similar Os.)

The size of the critical distance is determined by S's motivation to be correct in his recognition inference. If it is very important to be correct, the size of the critical distance will be very small and the nearest θ will have to be nearly identical to Ø to be accepted. On the other hand, if accuracy is less important, the critical distance will be large and the nearest θ can have cue values which are rather unlike Ø's and still be accepted.[3]

There is, in addition to the distance between the θs and the Ø, a second determinant of whether or not one of the θs will be accepted as having derived from the O in question. This is the frequency with which the θs previously have been experienced by S, i.e., their "familiarity" (Arnoult, 1956; Nobel, 1954). The θs which have occurred most frequently in the past can be expected to be more likely to appear in the future. Therefore, if one θ, or rather the O for which the θ stands, has been more frequently encountered than another, even though their distances from Ø are equal, S should be more inclined to accept the former as having derived from the present O. This influence of frequency of past encounters upon recognition is well documented by a sizable body of literature (e.g., Arnoult, 1956; Bruner, 1957; King-Ellison and Jenkins, 1954; Postman and Rosenzweig, 1956; Solomon and Postman, 1952).

If S decides that a particular θ is familiar

enough and that its cue values are similar enough to those which define the Ø, he can assume that they both derive from the O in question and the O is recognized. This clears the way for S to attribute to the O the class membership and all of the cue values which are associated with the Θ and to proceed with the interaction. If the O is not recognized, its class can be ascertained by either of two processes, assimilation or identification.

Assimilation. When the distance between the Ø and the nearest Θ in the cognitive space is greater than the crucial value, S can assume that they do not derive from the same O; however, the matter does not end there. If the Ø and Θ are reasonably similar, it is prudent to consider whether this similarity is not paralleled by a similarity in class membership and in as yet uninferred cue values. Thus, even though S knows that the O is probably not the same one which he has encountered before, he can still assume the Θ's class to be appropriate because the Θ and the Ø are so similar in other respects. When S is asked to give the appropriate class for an O which has never been encountered before but is very similar to a familiar O, the exhibited behavior is identical to that which would result from recognition (Beach, 1961, 1962). When S can clearly discriminate between the Θ and Ø and yet he adopts the Θ's class for the Ø, and therefore for the O, the process is called assimilation.

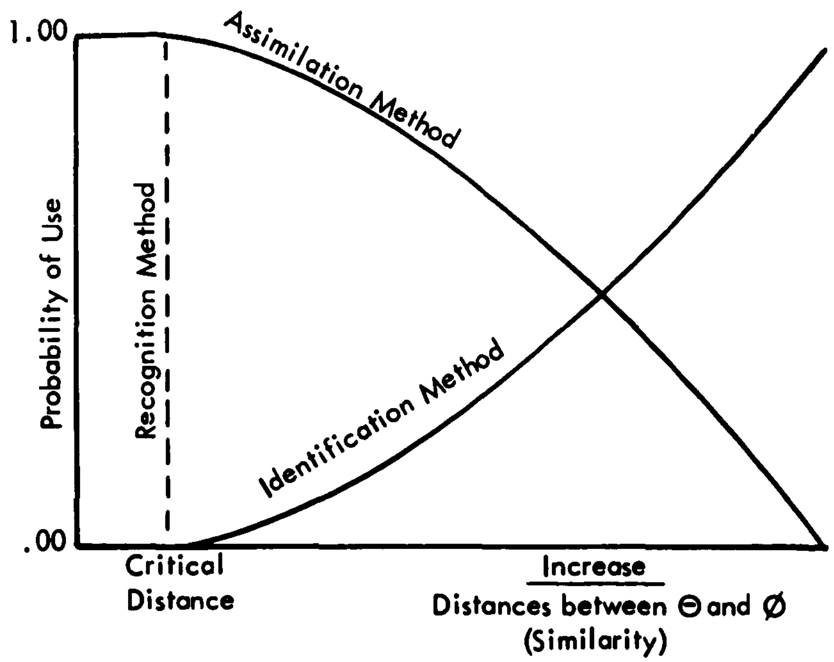

Fig. 2. Probability of use of each inference method as a function of the distance between Θ and Ø.

Figure 2

The hypothetical curves in Fig. 2 describe the effect of the distance between the Ø and the most similar Θ upon S's choice of an inference method. First, when familiarity is held high and constant, up to the critical distance, the Ø and Θ are assumed to derive from the same O and the present O is recognized as belonging to the class associated with the Θ. Just beyond the critical value, the Ø and Θ are still extremely similar and, even though he does not recognize the O, S is likely to go ahead and assimilate the Θ's class for the O. As the distance increases, the adoption of the most similar Θ's class for the O becomes less likely until at large distances it becomes extremely unlikely. At the same time, as the size of the distance increases S is more likely to obtain the O's class membership by the use of the identification procedure, which will be discussed presently.

In Fig. 3 are presented the hypothetical relationships between the frequency of past encounters with the $\underline{\Theta}$ and the probability of $\underline{S}$ using recognition, assimilation, or identification to make his inference. These two curves are from a family of curves which change as a function of the difference between $\underline{\emptyset}$ and $\underline{\Theta}$ defined in Fig. 2. (Placing the right margin of Fig. 3 against the left margin of Fig. 2 would produce a 3-dimensional graph which would define the family of curves.)

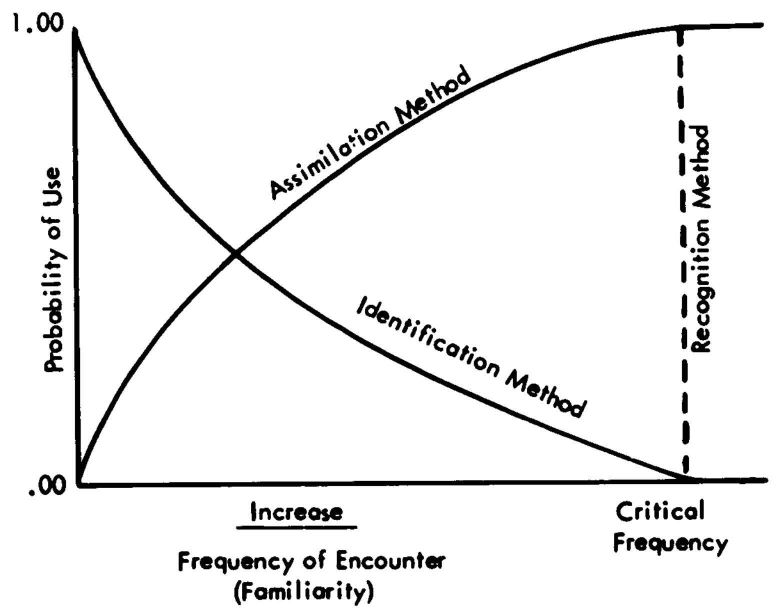

Fig. 3. Probability of use of each inference method as a function of the frequency of past encounters with the $\underline{\Theta}$.

Figure 3

When the distance between $\underline{\Theta}$ and $\underline{\emptyset}$ is held small and constant, an increase in the frequency with which $\underline{S}$ has experienced the $\underline{\Theta}$ results in an increase in the probability that $\underline{S}$ will use assimilation or recognition to infer the $\underline{O}$'s class membership and a decrease in the probability that he will use identification.

On the other hand, a larger distance between $\underline{\theta}$ and $\underline{\emptyset}$ results in a lower assimilation-recognition curve and a higher identification curve with corresponding changes in probabilities at any given frequency.[4]

It is important to note that there are as yet no data available to define the functions in Figs. 2 and 3 and that if the functions do exist, they probably vary in their exact shape and slope from one S to another as well as with an individual S's motivation to be correct in his classification of an O.

Identification. When an O's cue values define a $\underline{\emptyset}$ in the cognitive space which does not lead to recognition and which is too different from the nearest $\underline{\theta}$ for assimilation of the latter's class, S must obtain the class through identification. Unlike assimilation, identification is based upon a large number of $\underline{\theta}$s which are similar to the O, rather than upon just one. As every researcher knows, inferences based on generalizations derived from large samples have greater chances of being correct than those based on small samples. Similarly, when S's inferences for an unfamiliar O requires a high degree of accuracy, and when the $\underline{\emptyset}$ is too different from any $\underline{\theta}$ in the space, assimilation is abandoned and identification is used to infer O's class.

The identification process, like the recognition and assimilation processes, depends upon the immediately apparent cue values associated with the O to be identified. Each of these values has been experienced by the S in the course of his past experience with other Os, Os which were members of classes and which are represented by $\underline{\theta}$s in the cognitive space.

Given a specific level of the class hierarchy at which to work, each cue value associated with the O defines a relative frequency distribution on the classes. The distribution is composed of the $\underline{\theta}$s which possess the cue value, d, and which are members of each of the different classes, $\underline{k} = \underline{a}$ through $\underline{i}$. The proportion of $\underline{\theta}$s in each class, written $\underline{P(k/d)}$, will vary in size from one class to

another. If the past relative frequency of occurrence of events is assumed to be an indication of their future relative frequency of occurrence, the proportion for each class indicates the relative likelihood of the class being the one to which the O belongs. The best choice for the inference is the class with the largest P(k/d).

For one cue value this is a relatively straightforward process. When a number of the O's cue values are known, however, the decision becomes more complicated; the relative frequency distributions defined by the various cue values may not all indicate that the same class is the best choice for the O's class membership. Thus, in the situation in which he obtains the O's cue values on more than one cue dimension S must make a compromise among the indicated choices. This compromise can be made by using the proportions yielded by the cues for each class as weights which determine the expectation, E(k), that the class may be the best choice for the inference. These proportions, each of which constitutes $1/n^{th}$ of the information for each class, are indicative of the soundness of the assumption that the class in question is the best choice for S's inference about the O in view of his past experience with Os possessing the same cue values as the present O possesses.

The simplest way of reaching a compromise among the different cue values, and their associated relative frequency distributions, is to treat each of the cue values as independent factors influencing the choice for the inference. Then by summing the $\frac{P(k/d)}{n}$ over all cue values for each class one obtains an indication of the amount of "evidence" yielded by the n cue values for each of the classes, k = a through i.

$$E(k) = \sum_{d=1}^{n} \frac{P(k/d)}{n}$$

$$= \frac{P(k/1)}{n} + \frac{P(k/2)}{n} + \frac{P(k/3)}{n} + \ldots + \frac{P(k/n)}{n}$$

$$= \frac{P(k/1) + P(k/2) + P(k/3) + \ldots + P(k/n)}{n} \qquad (1)$$

k = a specific class under consideration as a possibility for O's identity. k=a through i.

d = a dimension on which the O's cue value is known. d = 1 through n.

Where E(k) is the total "evidence" indicating that each class, k, is the best choice for S's inference about the O's class membership. In spite of its rather unorthodox derivation, E(k) can be regarded as the likelihood that class k is the best choice for the inference, $\sum_{k=a}^{i} E(k) = 1.00$.

The simple additive form of equation (1) was deliberately selected because, while it is possible that other, more complex forms will prove to be more adequate, it is probably best to start with the simplest assumptions and then complicate the theory where need be. (Anderson, 1962, has successfully used a similar additive equation to predict Ss' impressions of Os based on descriptive adjectives.) Equation (1) has two particularly interesting properties. First, if any of the $\frac{P(k/d)}{n} = 0$, the equation does not reduce to E(k)=0 ~~as it~~ would if the equation were multiplicative. This seems to agree with common sense. Just because the cue value has never before been associated with a member of a particular class does not

mean that it is impossible that the present O belongs to the class; this may just be the first time S has encountered this situation. A second interesting property of the equation is that it is not necessary to weight the various $\frac{P(k/d)}{n}$ according to the ability of their corresponding cue dimensions to discriminate among the classes. Such a weighting procedure would be necessary if we used a correlational analysis (Hoffman, 1960; Todd, 1954) but is unnecessary here because the $\frac{P(k/d)}{n}$ will be nearly the same for every _k_ for a nondiscriminating cue dimension. For discriminating cue dimensions (i.e., highly correlated with the classes; e.g., Summers, 1962) the various values of $\frac{P(k/d)}{n}$ will be grossly different for different _k_s; this is, after all, what produces the correlation. However, the equation does leave open the possibility of adding weights to the $\frac{P(k/d)}{n}$ in order to take into account the salience of the various cue dimensions. In the present treatment we will ignore this possibility to simplify discussion.

An example may clarify how this equation can be used to identify _O_s. Let us assume that there are four classes of _O_s defined on the _S_'s cognitive space; _e_, _f_, _g_, and _h_. Next, assume an unfamiliar _O_ appears before _S_ and that on each of five dimensions it has the cue values; 1, 2, 3, 4, and 5. Each cue defines a distribution of _Θ_s on the classes; for cue value 1, .25 of the _Θ_s in the cognitive space are members of class _e_, .30 of them are in class _f_, .35 are in class _g_, and .10 are in class _h_. For cue value 2, .15 of the _Θ_s are in class _e_, .40 in class _f_, .25 in class _g_, and .20 in class _h_. Cue value 3 has a class distribution of .40 for _e_, .30 for _f_, .20 for _g_, and .10 for _h_. Cue value 4 has .20 for _e_, .25 for _f_, .30 for _g_, and .25 for _h_. And cue value 5 has .20 of its _Θ_s in class _e_, .50 in class _f_, .15 in class _g_, and .15 in class _h_. Substituting these proportions into the equation, we get:

$$E(e) = \frac{P(e/1)+P(e/2)+P(e/3)+P(e/4)+P(e/5)}{n}$$

$$= \frac{.25 + .15 + .40 + .20 + .20}{5}$$

$$= \frac{1.20}{5}$$

$$= .24$$

$$E(f) = \frac{P(f/1)+P(f/2)+P(f/3)+P(f/4)+P(f/5)}{n}$$

$$= \frac{.30 + .40 + .30 + .25 + .50}{5}$$

$$= \frac{1.75}{5}$$

$$= .35$$

$$E(g) = \frac{P(g/1+P(g/2)+P(g/3)+P(g/4)+P(g/5)}{n}$$

$$= \frac{.35 + .25 + .20 + .30 + .15}{5}$$

$$= \frac{1.25}{5}$$

$$= .25$$

$$E(h) = \frac{P(h/1)+P(h/2)+P(h/3)+P(h/4)+P(h/5)}{n}$$

$$= \frac{.10 + .20 + .10 + .25 + .15}{5}$$

$$= \frac{.80}{5}$$

$$= .16$$

The $\underline{O}$ possessing these cue values therefore should be identified as a member of class $\underline{f}$, the class with the highest value for the equation. When the $\underline{O}$'s class has been inferred, the $\underline{\emptyset}$ takes on the class as an attribute and becomes a permanent $\underline{\theta}$ in the cognitive space.

Relevant Research. The hypothesized nature of the class inference methods, as well as the contents of Fig. 2, is based upon the results of recent research by Beach (1961, 1962). In these studies large sets of cards served as populations of $\underline{O}$s. Each card had on its face a cue value from each of three dimensions (five consecutive values on a letter dimension, number dimension, and a pointer position dimension in one experiment and twelve consecutive values on each dimension in the other experiment.) Written on the back of each card was the name of the class to which it belonged (equal numbers of cards belonged to each of the classes, Red, Yellow, or Blue in one experiment and to each of the classes, Red, Yellow, Blue, Green, or Orange in the other). The cue values on the cards were selected so that associated with each was a relative frequency distribution similar to that in the example given above. This permitted the application of equation (1) to each card and the derivation of an $\underline{E(k)}$ for each possible class. The class for which the $\underline{E(k)}$ was highest was assumed to be the class to which $\underline{S}$ should assign the card if the theory was correct.

To test the applicability of the theory some of the cards were deliberately assigned to classes which did not have the highest $\underline{E(k)}$. If the theory could predict inferences correctly, it was expected that $\underline{S}$s' inferences for these "trick" cards would be the class with the highest $\underline{E(k)}$ from equation (1) and, as a result, consistently would be in error. In addition to the trick cards some test cards were constructed to be shown to the $\underline{S}$s at the end of training.

The training procedure consisted of showing the Ss each card in the set (120 cards for one experiment and 200 cards for the other) and having them infer to which class each card belonged. After the inference was made the experimenter read the correct answer from the back of the card. This procedure was repeated for a number of training sessions and then the Ss were shown the test cards. By using this method, data were obtained for repeated, corrected inferences (for the trick and nontrick cards in the training set) and for one trial, noncorrected inferences (for the test cards).

The results for both experiments were that, for nontrick cards, i.e., those for which the high theoretically derived E(k) corresponded to the correct answer, Ss' inferences were highly accurate. For the trick cards, however, Ss tended to give the high E(k) class as their inference and thereby to be incorrect. This tendency for the trick cards was particularly strong in the early stages of training but as training progressed about as many Ss gave the correct class as their inference as gave the high E(k) class. This result implies that as training progressed, and the frequency of contact with the cards increased, some other mechanism took over and allowed the Ss to begin to detect the trick cards and to give the correct answers.

For each of the test cards equation (1) produced an E(k) for each possible class. As with the cards in the training set, it was predicted that the Ss would assign the test cards to the classes having the highest E(k). This was found to be true only when the test card was markedly dissimilar from all of the cards in the training set. When the test card's cue values were similar to one of the previously experienced card's, the Ss' inference for the test card's class was the class associated with the card (either trick or nontrick) from the training set.

These results have had a strong influence on the form of the present theory. In the first place

the test card results indicate that Ss depend heavily upon the similarity between a previously experienced O and the one for which the inference is to be made in order to reach a decision about the latter's class. The trick card results imply, however, that before recognition or assimilation can take place it is necessary for the old O to have been experienced a number of times (or perhaps for a fairly long stretch at one time). The latter conclusion is based upon the fact that Ss utilized the identification method in the early stages of training and only came to depend upon the alternative method, which we assume to be recognition of the individual cards each time they appeared, after having seen the cards a number of times.

It is assumed, therefore, that Ss first attempt to infer an O's class by the recognition method, accepting as possible candidates only those Θs which have been repeated experienced. If no repeatedly experienced Θ is sufficiently similar to the Ø defined by the O's cue values, S resorts to assimilation--still requiring that a repeatedly experienced Θ be the basis of the assimilation. If no Θs are sufficiently similar to the O in question, S turns to the identification method, a method which apparently utilizes every Θ in the cognitive space rather than just the most frequently experienced. The results of these experiments will be discussed further in relation to inferences about O's cue values.

Inference of Cue Values

When an O is encountered, it can be considered the possessor of two kinds of cue values, known and unknown. The known cues include all of the immediately apparent aspects of the O which are used to make inferences at the lower levels of the cue dimension hierarchy and which form the basis for the inference about the O's class. After the class is inferred all of the cue values associated with the O are considered known cue values and all other cue values are considered unknown cue values. Chief

among the unknown cues are the higher-order cue values upon which S̲s' interactions with O̲s are based, e.g., intelligence, distance, size, and friendliness. These higher-order cue values, more so than an O̲'s class, are the end point in S̲'s search for knowledge about the O̲. It is only by knowing these attributes of the O̲ that S̲ can properly plan his interaction and attain his long range goals.

While many cue inferences have taken place before S̲ infers an O̲'s class, i.e., at the lowest levels of the cue dimension hierarchy, let us first discuss those cue inferences which are made after the class has been inferred. This order of discussion will simplify the explanation of the mechanics of the lower-order, pre-class, inferences.

Recognition and Cue Values. Just as recognition of an O̲ reveals its class membership, so too does it reveal many of the O̲'s heretofore unknown cue values. When the O̲ is first encountered, the only attributes it exhibits are lower-order cue values. When these cues permit recognition of the O̲, much more is immediately known about it; all of the cue values and the class associated with the Θ̲ in the cognitive space can be attributed to the O̲. These cue values and class have been gleaned through past experience with the O̲ and are retrieved from the cognitive space through knowledge of the Θ̲'s location in the space. In this way recognition permits the utilization of past experience in the planning and execution of interaction with the O̲.[5]

The cue values associated with the Θ̲ are, however, only those which past experienced has provided. If other, as yet unknown, cue values are required for the interaction, S̲ must fall back on either assimilation or identification in order to infer them.

Assimilation and Cue Values. When an O̲ is not recognized, or when specific attributes of a recognized O̲ have never been experienced, S̲ can utilize assimilation to provide the unknown cue values. The procedure is essentially the same as for assimila-

tion of classes; the needed cue values can be assimilated from the most similar $\underline{O}$ in the cognitive space. There is one difference between class and cue assimilation, however. Class assimilation is usually based on only one $\underline{O}$ but it is possible that cue value assimilation may be based on a number of $\underline{O}$s. This is because the most similar $\underline{O}$ may not possess a cue value on the cue dimension needed for the $\underline{O}$ in question. If, however, a somewhat less similar $\underline{O}$ possesses a value on the dimension, it may be assimilated. When a number of cue values are needed, it is possible that they could all derive from different $\underline{O}$s in the space. Of course, the hypothetical function in Fig. 2 would still govern whether or not any particular $\underline{O}$ was too dissimilar to be utilized for assimilation of a given cue and, in addition, $\underline{S}$'s motivation to be correct, governed by the importance of the assimilated cue value to the coming interaction, as well as the frequency with which the $\underline{O}$ has been previously experienced, would also influence the acceptance or rejection of the $\underline{O}$ for cue value assimilation.

Identification and Cue Values. When neither recognition nor assimilation provides the required cue values for an $\underline{O}$, they can still be inferred in much the same way that classes are identified. When the cue inference is begun, the $\underline{O}$'s class is already known through recognition, assimilation, or identification. Moreover, some of its lower-order cue values are known and sometimes, if it has been recognized or if cue values from other $\underline{O}$s have been assimilated, some of its higher-order cue values are also known. Knowledge of the class together with each of the known cue values provides the foundation for inferring the necessary cue values.

Knowledge of an $\underline{O}$'s class is seldom an end in itself. The class is super-ordinate to the cues and its value lies in the fact that, by and large, the $\underline{O}$s which are its members are more similar to each other than they are to the rest of the $\underline{O}$s in

the space. This fact makes knowledge of an *O*'s class important, it defines a group of *O*s upon which the inferences about an *O*'s unknown cue values can be based (Bolles and Bailey, 1956). However, these inferences are dependent not only upon knowledge of what the *O* is, its class, but also upon knowledge of its particular attributes, its known cue values. Each of the known cue values defines a subclass within the larger class, a subclass composed of class members which possess that specific cue value. It is these subclasses, one for each of the *O*'s cue values, upon which the inferences are based.

In terms of the cognitive space, the *O* possesses both known cue values and a known class; the goal is to infer a cue value on a specific cue dimension for which the value is unknown. The reasoning underlying this inference is parallel to that upon which the class identification method is based. Here, however, *S* need not deal with all of the *O*s in his cognitive space. He can begin by considering only the *O*s that belong to the *O*'s class as relevant to the inference; if the *O* is a chair, only previously experienced chairs need be considered in determining the inference. Then each of the *O*'s cue values defines a relative frequency distribution upon the unknown cue dimension. This distribution yields the proportion of *O*s (1) that belong to *O*'s class, *k*, (2) that possess the same cue value, as the *O* possesses, *d*, and (3) that possess each of the possible values of the unknown cue dimension, *c* = *1* through *j*. This proportion, *P(c/k,d)*, indicates the likelihood that a value, *c*, on the unknown cue dimension is the best choice for the inference given the *O*'s class and its value on one of the known cue dimensions. As in the inference of *O*'s class, the cue value associated with the highest proportion is the best choice for the inference about the *O*'s unknown cue value.

When, as is usually the case, *S* knows more than one of the *O*'s cue values there may be conflicting inferences yielded by the different dis-

tributions. As in the inference of classes, a compromise can be made among these conflicting inferences by summing, for each unknown cue value, the proportions contributed by each known cue value (keeping in mind that each proportion is only $1/n^{th}$ of the evidence for that specific value of the unknown cue dimension).

$$E(c) = \sum_{d=1}^{n} \frac{P(c/k,d)}{n}$$

$$= \frac{P(c/k,1)}{n}+\frac{P(c/k,2)}{n}+\frac{P(c/k,3)}{n}+\ldots+\frac{P(c/k,n)}{n}$$

$$= \frac{P(c/k,1)+P(c/k,2)+P(c/k,3)+\ldots+P(c/k,n)}{n} \qquad (2a)$$

c = a cue value under consideration as O's unknown cue. c = 1 thru j.
d = a dimension upon which the O's cue value is known. d = 1 thru n.
k = O's class.
n = the total number of dimensions upon which O's cue values are known.

where $E(c)$ is the total "evidence" indicating that each cue value, c, from the unknown cue dimension is the best choice for S's inference.

As with $E(k)$ from equation (1), $E(c)$ can be regarded as the likelihood that cue value, c, is the best choice for the inference, $\sum_{c=1}^{j} E(c) = 1.00$, and S's inference should be the cue value associated with the highest value of $E(c)$.

LOWER-LEVEL CUE INFERENCES. For inferences of the basic, lower-level cue values, those inferred from sensation and used to infer classes, equation (2a) must be modified. At this stage S knows only the very basic sensation cues--which are multidimensional at the level of the nerve if not at the level of the neuron--and perhaps a few other basic

cue values which have already been inferred. The class is not known and therefore it must be dropped from equation (2a), which becomes:

$$E(c) = \sum_{d=1}^{n} \frac{P(c/d)}{n}$$

$$= \frac{P(c/1)}{n} + \frac{P(c/2)}{n} + \frac{P(c/3)}{n} + \ldots + \frac{P(c/n)}{n}$$

$$= \frac{P(c/1) + P(c/2) + P(c/3) + \ldots + P(c/n)}{n} \qquad (2b)$$

where each proportion consists of those $\underline{\Theta}$s which possess both the known cue value, $\underline{d}$, and each cue value, $\underline{c}$, from the dimension being inferred. As in equation (2a), the cue value with the highest $\underline{E(c)}$ is the best choice for $\underline{S}$'s inference of $\underline{O}$'s unknown cue value.

Through the inference process described by equation (2b) the $\underline{O}$'s basic cue values are built up and stored in the cognitive space in terms of the location of the $\underline{\emptyset}$. After enough values are known, or when no more are forthcoming, the $\underline{O}$'s class is determined through recognition, assimilation, or identification. Then, higher-order cues can be inferred through their association with the recognized $\underline{\Theta}$, through their assimilation from a similar $\underline{\Theta}$, or, now that the class is known, through identification inferences and equation (2a).

Relevant Research. When an $\underline{O}$ is recognized, its cue values are immediately revealed for a large number of cue dimensions. This is a familiar and obvious happening which is exactly what we mean in the everyday use of the word "recognize". For assimilation and identification, however, the cue inference process is less apparent. These require experimental investigation in order to test the adequacy of the present theoretical formulation. As yet, little appropriate research has been done.

Some light is cast upon the assimilation and iden-

tification processes by data from the previously described study by Beach (1961). In this study, you will remember, Ss repeatedly made inferences about the class memberships for each card in a large training set. Then they were shown a number of test cards, which they had never seen before, for which they also made inferences. The aspect of the study which is of interest here concerned S's inferences about the values for cues which were missing from some of the test cards. In most cases the inferences about the unknown cue values were dictated by assimilation; the cue value possessed by a training card which possessed cues similar to the test card's known cues was given as the test card's missing cue. In some cases, however, identification played a role and the inference was predicted by equation (2a). Which method best predicted the Ss' inference for a given test card appeared to be governed by the degree of dissimilarity between the test card and the most similar card from the training set. While the results are not sufficient to yield specific information about how the judgments were made, they do show that assimilation and identification play a role in Ss' inferences of unknown cue values.

Aside from the results just cited--which are themselves much too vague to be seriously considered as evidence for the theory--there are two studies which are relevant to the role of the identification process in cue inference. These studies indicate the importance of knowledge of the O's class in the inference of unknown cue values. In the first of these (Goodnow, 1954), which is comparable to the study cited above, Ss were taught to classify Os on the basis of three dichotomous cue dimensions. Then they were shown test Os, each of which had one missing cue, and asked to infer the missing value. It was found that the two known cues apparently were used to determine the Os' classes and then the inferences of the missing cues were based on the classes. Unfortunately, the results do not yield information about whether or not the Os' two known cues

influenced the inference in the manner described by equation (2a) or whether the inferences were determined by the class alone. Nevertheless, these results indicate that inference of the O's class is a first step in the inference of unknown cue values and that these unknown cue inferences are influenced by S's knowledge of the O's class membership.

The second relevant study is by Bolles and Bailey (1956). Here five Ss were told 54 Os' class memberships together with some of their cue values (e.g., "I have an ashtray, it is a ceramic ashtray with V-shaped notches in it.") and asked to infer the Os' sizes. The average correlation between the inferences and the Os' actual sizes was .994. In addition it was found that Ss' errors were larger for the classes of Os with which they had had less experience and for the classes which had large size variance among their members. The authors interpret these results as showing the importance of prior knowledge of the O's class upon unknown cue inferences. The additional cues given about the Os are confounded, however, with the knowledge of the class, and it seems safer to conclude that the accuracy is the result of knowing both the class and the additional cues. Consequently, the relative roles of these two factors are not known and it is impossible to tell whether or not equation (2a) could account for the results.

These three studies are not a very impressive display of support for the theory. At present the author is undertaking research which will attempt to examine the cue inference processes more thoroughly.

INFERENTIAL ACCURACY

The Revision Process

As we have seen, inferences about an O's attributes, its class and its cue values, can be made

by any of three methods: recognition, assimilation, or identification. Whichever method is used, however, the S needs his inferences to be accurate, first, because they must stand up when they are used in interactions with the O and, second, because they become associated with the Θ in the cognitive space and form the basis for future interactions when the O is encountered again. To assure the utmost accuracy for the class and cue values associated with a Θ each inference is constantly being re-evaluated in light of information gained in the course of interaction with the O.

The revision process consists of re-inferring each attribute of the Θ by the identification method and of comparing the new inference with the old one, no matter which method originally produced the latter. If the old and the new inferences agree, it is an indication of the stability of the Θ's attribute and of the accuracy of the previous inference. When such agreement occurs, the inference is retained as an attribute of the O and the fact that it was affirmed is recorded for future reference. In terms of the cognitive space, recording a cue value's affirmation is accomplished by "frequency of affirmation dimensions" which are associated with each Θ in the space and which represent the array of classes and each of the cue dimensions. Upon each affirmation dimension is tallied the frequency with which each class and each value on the corresponding cue dimension has been experienced as an attribute of the Θ.

Affirmation frequency and inference. When an O is encountered for which a Θ exists in the cognitive space, the Θ's affirmation distributions are used to decide about any of the O's attributes for which, on that encounter, no information is immediately available. This would be a very straightforward process if S merely assigned the attribute value (class or cue value) most recently associated with the Θ. It is probably not this simple however. If it were, S would completely change his estimation

of the O every time it evidenced even a temporary change in any of its attributes. (If, for example, S's wife, who is usually pleasant, was particularly short tempered yesterday, he would have to change completely his entire estimation of her disposition. This, of course, would be quite an unreasonable thing for him to do.) On the other hand, S could always assume the attribute value which has been most frequently associated with the Θ to be the best inference each time the O is encountered. While this is probably a better strategy than always assigning the most recent value of the attribute, it still is not satisfactory. Always assigning the most frequent value of the attribute completely ignores all changes in the O until one of the other values occurs so frequently as to outweigh the old most frequent value. It would appear, therefore, that some compromise is in order; both the most frequently occurring value of the attribute (the mode of the affirmation distribution) and the most recently occurring value of the attribute should jointly determine the selection of the value which is to be assigned to the Θ. While the exact formulation for the determination of the compromise has yet to be done, the compromise probably should lie somewhere between the mode of the affirmation distribution and the value of the attribute which has been most recently associated with the O. Due to the mode's higher relative frequency of occurrence, the compromise should, perhaps, be somewhat nearer to the mode than to the most recent class or cue value.[6]

The concept of affirmation frequency is vital to the theory because it is used every time S "recalls" the cue values associated with a recognized or assimilated O. In addition, as the following discussion will reveal, it serves to tie together many apparently unrelated parts of the theory.

Affirmation frequency, recognition, and assimilation. In the discussion of the factors which influence the selection of a Θ for recognition or assimilation it was stated that one important determinant was the frequency with which S had previously

experienced the $\underline{\Theta}$ (cf., Fig. 3). The affirmation dimension for the classes provides the method by which this experiential frequency is recorded in the cognitive space. Because every $\underline{\Theta}$ is classified on each encounter the sum of the frequencies on the affirmation dimension for classes indicates the frequency of $\underline{S}$'s encounters with the corresponding environmental $\underline{O}$. Of course, whether or not this sum, together with the distance between the $\underline{\Theta}$ and the $\underline{\emptyset}$, is sufficient for $\underline{S}$ to utilize $\underline{\Theta}$ for recognition or affirmation is dictated by the $\underline{S}$'s motivational level at the moment. This matter was previously discussed in some detail and no further comment is necessary here.

Affirmation frequency and revision. When a $\underline{\Theta}$'s attribute is re-evaluated and affirmed, it is retained as one of the classes or cue values and its frequency of affirmation is increased. In this case there is no revision of the $\underline{\Theta}$'s attribute; indeed, the existing attribute is strengthened as a part of the $\underline{\Theta}$.

If, however, the new inference disagrees with the old one, a decision must be made about which of the two is correct. This decision is made by comparing two factors, the frequency with which the old inference has been affirmed and the magnitude of the $\underline{E(k)}$ or $\underline{E(c)}$ for the new inference. Obviously, if the old inference has been frequently affirmed, it is not prudent to replace it with the new inference unless the latter is shown to be very probable in light of new information about the $\underline{\Theta}$. On the other hand, an old value which has been affirmed only a few times is not necessarily a better choice than the new inference.

Inferential confidence. To compare frequency of affirmation and the magnitude of $\underline{E(k)}$ or $\underline{E(c)}$ it is necessary to introduce another concept into which they can both be transformed. This concept is degree of inferential confidence. It is positively related to both frequency of affirmation and to magnitude of $\underline{E(k)}$ and $\underline{E(c)}$. Thus, to decide which of the two inferences to assign to the $\underline{O}$, the $\underline{S}$ must compare the confidence value corresponding to the old inference, determined

by frequency of affirmation, to that of the new inference, determined by magnitude of E(k) or E(c). The inference with the higher confidence value is the one assumed to be correct for the O.

The concept of inferential confidence ties together many of the mechanisms which have been introduced throughout the previous discussion. It is defined as the degree to which S believes that a decision or an inference is accurate and it can be measured in a number of ways, although each method has its drawbacks. The most candid method, and one with apparently high validity, is merely to ask S how confident he feels about a decision or inference (see, for example, Irwin, 1953; Pollack and Decker, 1958; Anderson and Walen, 1960). The drawback here is that the response scale chosen by S may not relate linearly to confidence. A more complex method is to ask S to place bets on his decisions or inferences (Edwards, 1961); however, this method is complicated by the value of the reward. Another method, which to the author's knowledge has never been used, is to utilize the degree of specificity or generality in S's interactions with an O to determine his confidence in the inferences upon which the interaction is based. If S is confident about the accuracy of his inferences about an O, those aspects of the interaction which are based on the inferred attributes should reflect the confidence. In this situation, high inferential confidence should be revealed in the form of fairly specific predictions about the O and the O's reaction to S's behavior toward it. Because S knows about this particular aspect of O he can confidently and accurately predict the course of the interaction. If, on the other hand, S lacks confidence in his inferences about the relevant attributes of the O, he can make more cautious ("cagey") predictions about O. Less specific, general, predictions will permit interaction--and thus new knowledge with which to re-evaluate the questionable inferences--without jeopardizing S's goals through a *faux pas*. This sort of behavior is frequent in social situations in which S's interaction with the O is dictated solely by social form until he has an opportunity to "size-up" the O. Proper exami-

nation of the generality and specificity of behavioral interactions with Os could possibly lead to a method of estimating S's confidence in the inferences underlying the interaction.

Inferential confidence is essentially a scale of subjective probabilities about the correctness of an inference. Each inference is associated with a value from this scale, and the critical value for acceptance or rejection of the inference is determined by S's motivation to be correct. Thus, for example, in Fig. 2, the critical distance is the degree of difference between Θ and $\emptyset$ beyond which S's inferential confidence is too low to accept that they derive from the same O. Similarly, in the same figure, the rejection of assimilation and the switch to identification is governed by S's confidence about assimilation inferences when large distances exist between the Θ and $\emptyset$ or when the frequency of contact with the Θ is small.

Inferential confidence also plays an important role in the selection of the level in the class hierarchy where an O's class will be inferred. If the E(k) for the most likely class on one level is not large enough to warrant sufficient inferential confidence, S can move up the hierarchy until a level is found which possesses a class with a satisfactory E(k). While S's motivation to make a correct class inference will tend to make him select a high level in the hierarchy at which to work, his motivation to be correct on subsequent cue inferences will tend to make him select a low level. The hierarchy level at which the class inference is finally made will be where S's degree of confidence for the class inference balances the degree of confidence for subsequent cue inferences.

As these examples indicate, inferential confidence provides a common link for many of the concepts introduced throughout the theory. Through this concept motivation can be linked to the choice of inference method, to probabilities yielded by the identification equations, to frequency of inference affirmation, and to the selection of the level in the

class hierarchy at which a class inference will be made. In addition to all of this, however, inferential confidence also governs another aspect of S's inference behavior--the search for new information about the O.

Cue search. When inferential confidence is below the degree required by the existing level of motivation, the inference is assumed to be of questionable validity. When this happens some sort of action is required to remedy the situation. The action usually consists of a search for further information upon which the inference can be re-evaluated and either affirmed or revised. This search for information, called cue search (Bruner, Goodnow, and Austin, 1956), may be either active probing of the environment through questions, manipulation, reading, or the like, or it may be passive reception of information in the course of cautious interaction with the O. In either case, it consists of behavior in which S obtains totally new information about an O or in which he receives information he already possessed. The former is used to expand his knowledge about the O and to re-evaluate, and perhaps revise, previous inferences. The latter adds to the various affirmation frequencies and, consequently, raises S's inferential confidence about those particular attributes of the O. If S is not wholly confident that he recognizes an O he can search for information about the O. In this situation cue search will be initiated when the distance between the Θ and the Ø is very near the crucial distance. If the O is indeed familiar to S, the additional information will usually decrease the size of the distance and thereby permit recognition. If it is unfamiliar, the distance will usually increase and the O will not be recognized.

Inferences based upon identification will lead to cue search if no E(k) or E(c) is sufficiently high to warrant confidence in the accuracy of the inference. Indeed, whenever a number of inference alternatives are approximately equally likely, whether they are

based on recognition, assimilation, or identification, cue search will usually be undertaken to break the deadlock.

The long run importance of cue search, and the revision which results from it, is to provide S with accurate and reliable knowledge about the Os with which he interacts. This not only assures more successful interactions with these Os but it also provides a solid foundation of knowledge for inferences about Os which are encountered for the first time. Although Irwin, Smith, and Mayfield (1956) found that Ss' inferential confidence was greater when they possessed more information about an event than when they possessed less, a quantitative statement of the relationship is not available. It is clear, however, that because they are measurable, these two concepts are important keys to the investigation of the remainder of the theory. By appropriate manipulation of motivation and opportunity for cue search, as well as by the proper selection of cue values associated with experimentally controlled Os, it should be possible to learn about the empirical relationships which are assumed to exist between these and other theoretical concepts.

GENERAL REMARKS

The theory is an attempt to provide a framework within which various aspects of behavior can be examined and interrelated. Specifically, it is an attempt to delineate the role of past experience with Os in the determination of inferences on subsequent encounters with the same or different Os. The term "past experience" has always been psychologists' refuge from ignorance. When called upon to explain why Ss' behavior fails to demonstrate a 1:1 relationship with complex stimuli, we glibly attribute it to the influence of "past experience." As the Gestaltists pointed out at length, the indiscriminate use of this explanation amounts to no explanation at all. It is not until we can begin to specify the manner in which past experience influences responses to complex stimuli that the term comes to have much scientific value.

In the present theoretical treatment of the effects of past experience upon inference behavior, we have chosen to view stimuli as conglomerates of cues. For recognition and assimilation, nonlinearity between the _O_'s immediately apparent cue values and the _Ss_' behavior is attributed to _Ss_' additional knowledge about the _O_, knowledge retained by the location of a _Θ_ in the cognitive space. For identification the theory is a bit more complex. Here it is necessssary to assume a linear relationship between _each_ of the _O_'s immediately apparent cue values and the possible inference alternatives, i.e., between each $P(k/d)$ and $E(k)$ or each $P(c/k,d)$ and $E(c)$, then nonlinearity in the responses to the complex of immediately apparent cues which comprises the _O_ is attributed to the combination of all of these linear inferences in the manner described by equations (1), (2a) or (2b).

The theory is a decision theory, but it is more. It is an attempt to place inference behavior within a framework in which a number of factors ("subjective probabilities," motivation, inferential confidence, past experience with _O_s, etc.) can interplay. The aim is to produce an experimentally testable theory which will lead to a greater understanding of inference, its laws, and its relationships to other areas of psychological functioning. Equally important, an attempt has been made to make the general approach of the theory, although not necessarily its specific mechanisms, conform to a common sense view of how inference behavior takes place.

The specific roles of recognition, assimilation, and identification in the attaining of knowledge about _O_s, and the effects of these inference methods upon subsequent inferences, fit well with common sense notions about cognitive functioning. Indeed, recognition _is_ the equating of a present _O_ with a previously experienced one, and the common behavioral pattern is to utilize previously obtained knowledge to interact with the _O_. This is precisely what we mean when we say we recognize our car ("the carburetor is out of whack"), or our boss ("he prefers to be called J.R."), or our house ("I can walk in without ringing the

bell"). The process which we have called assimilation is also commonly experienced. The techniques used to operate one typewriter work for a similar one even though the two are not identical; behavior toward a person perceived as highly similar to a person whom you know well is a cautious version of the behavior which would be appropriate to the latter, etc. The identification process takes care of the left-over Os, those which are too dissimilar to any Os to be either recognized or assimilated. Whether or not the specific mechanisms proposed in the theory are entirely correct, these three kinds of inferences make sense and apparently cover the field.

The cue and class hierarchies are also common sense. When it is realized that Os must be, so to speak, "reconstructed" from sense data, the hierarchical structure of the cues becomes a necessity. It is clear that there are very real differences among cues; some are closely akin to sensory events (intensity of sounds, retinal contour, temperature), others are more abstract (friendliness, three-dimensional shape, senility). Granted that they all derive from sensation rather than from innate ideas, the abstract cues must be generated by the sensory events and they in turn generate other knowledge about the Os. It is equally clear that what an O *is* influences our interpretation of its properties; therefore, the O's class must be known before many of its attributes can be properly evaluated.

The class hierarchy is common sense if we realize that every class of Os can be subdivided by segregating its members according to specified similarities among their properties and giving the resulting subgroups new generic names. Assuming that all classes result from this process of differentiation and renaming, it is easy to see how a class hierarchy comes to exist as S has increased contact with Os and is required to be increasingly discriminate and correct in his interactions with them. For example, the broad class "aircraft" becomes increasingly differentiated by an aviator as he gains more and more experience with them. "Aircraft" become either "jets"

or "props," and the "props" are differentiated into "reciprocating engines" and "turboprops." These two classes are further subdivided into smaller classes which are also subdivided, etc. The result is a hierarchy which extends from "aircraft" (which is itself subsumed under some higher-order class) down to the particular planes with which the aviator has had personal experience.

Another common aspect of our experience with Os is their apparent individuality. Even though they are usually similar to other Os, we experience them as unique individuals which differ, no matter how slightly, from their fellows. In psychological theorizing, on the other hand, there is a tendency to forget this fact. In the present theory an effort has been made to acknowledge and to utilize an O's individuality as an important part of the inference process. In the first place it is assumed that Ss' inferences are normally made for individual Os and are unique for that O rather than for runs or highly repetitive occurrences of the O. Each inference is ordinarily aimed toward maximum accuracy for the effort involved; if all else is equal and increased accuracy is available only with strenuous cue search, S will settle for less accuracy.

An O's individuality is also reflected in recognition and in its temporal priority in the class inference process. Recognition is the quintessence of dependence upon an O's individuality and constancy over time. Similarly, cue inference through the identification process emphasizes the O's individuality by stressing the role played both by the O's class and by its unique cue values. The O is not merely relegated to a class and, once there, merged into the mass of Os which are its members. Rather, inferences about its cues are based on only those fellow class members which share its unique cue values.

And, finally, the integrity of the individual O is reflected in the theory's emphasis on affirmation

and revision of knowledge about each of them. In this way the onward flow of interaction with an _O_ is aided as the _O_'s eccentricities and peculiarities become known and as old incorrect knowledge is revised.

REFERENCES

Ames, A. Nature and origin of perceptions (preliminary laboratory manual). Hanover, N. H.: The Hanover Institute, 1949. Cited in C. E. Osgood, Method and theory in experimental psychology. N. Y.: Oxford Univ. Press, 1956.

Ames, A. An interpretive manual: The nature of our perceptions, prehensions and behavior. Princeton, N. J.: Princeton Univ. Press, 1955.

Anderson, N. H. Application of an additive model to impression formation. Science, 1962, 138, 817-818.

Anderson, N. H., and Walen, R. Likelihood judgments and sequential effects in a two choice probability learning situation. J. exp. Psychol., 1960, 60, 111-120.

Arnoult, M. D. Familiarity and recognition of nonsense shapes. J. exp. Psychol., 1956, 51, 269-276.

Attneave, F. Psychological probability as a function of experience frequency. J. exp. Psychol., 1953, 46, 81-86.

Beach, L. R. A mathematical theory of inference behavior. Unpublished doctoral dissertation, Univ. of Colorado, 1961.

Beach, L. R. A study of inference behavior. Project MR005.13-5001 Subtask 12, Report No. 2. Pensacola, Fla.: U.S. Naval School of Aviation Medicine, 1962.

Binder, A., and Feldman, S.E. The effects of experimentally controlled experience upon recognition responses. Psychol. Monogr. 1960, 74, 9.

Bolles, R. C., and Bailey, D. C. Importance of object recognition in size constancy. J. exp. Psychol., 1956, 51, 222-225.

Boring, E. G. Sensation and perception in the history of experimental psychology. New York: Appleton-Century, 1942. Cited in J. S. Bruner, Jackqueline Goodnow, and G. A. Austin, A study of thinking. New York: Wiley, 1956.

Bruner, J. S. On perceptual readiness. Pschol. Rev., 1957, 64, 123-152.

Bruner, J. S., Goodnow, Jacqueline, and Austin, G. A. A study of thinking. New York: Wiley, 1956

Bruner, J. S., and Minturn, A. L. Perceptual identification and perceptual organization. J. gen. Psychol., 1955, 53, 21-28.

Bruner, J. S., Postman, L., and Rodrigues, J. Expectation and the perception of color. Amer. J. Psychol., 1951, 64, 216-227.

Brunswik, E. Perception and the representative design of experiments. Berkeley: Univ. of Calif. Press, 1956.

Cronback, L. J., and Gleser, Goldine. Assessing similarity between profiles. Psychol. Bull., 1953, 50, 456-473.

Edwards, W. Behavioral decision theory. Annu. Rev. Psychol., 1961, 12, 473-498.

Ekman, G., Goude, G., and Waern, Y. Subjective similarity in two perceptual continua. J. exp. Psychol., 1961, 61, 222-227.

Gardner, R. A. Probability learning with two and three choices. Amer. J. Psychol., 1957, 70, 174-185.

Goodnow, R. E. Utilization of partially valid cues in perceptual identification. Unpublished doctoral dissertation, Harvard Univ., 1954.

Hake, H. W., and Hyman, R. Perception of the statistical structure of a random series of binary symbols. J. exp. Psychol., 1953, 45, 64-74.

Hoffman, P. J. The paramorphic representation of clinical judgment. Psychol. Bull., 1960, 57, 116-131.

Irwin, F. W. Stated expectations as functions of probability and desirability of outcomes. J. Pers., 1953, 21, 329-333.

Irwin, F. W., Smith, W. A. S., and Mayfield, Jane F. Tests of two theories of decision in an "expanded judgment" situation. J. exp. Psychol., 1956, 51, 261-265.

Ittleson, W. H., and Slack, C. W. The perception of persons as visual objects. In R. Tagiuri and L. Petrullo (Eds.), Person perception and interpersonal behavior. Stanford: Stanford Univ. Press, 1958.

Kelly, G. A. The psychology of personal constructs. Vol. I. New York: Norton, 1955.

King-Ellison, Patricia, and Jenkins, J. J. The durational threshold of visual recognition as a function of word-frequency. Amer. J. Psychol., 1954, 67, 700-703.

Miller, N. E., and Dollard, J. Social learning and imitation. New Haven: Yale Univ. Press 1941.

Nobel, C. E. The familiarity-frequency relationship. J. exp. Psychol., 1954, 47, 13-16.

Osgood, C. E., and Suci, G. J. A measure of relationship determined by both mean difference and profile information. Psychol. Bull., 1952, 49, 251-262.

Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. The measurement of meaning. Urbana, Ill.: Univ. of Ill. Press, 1957.

Overall, J. E., and Williams, C. M. Models for medical diagnosis. Behav. Sci., 1961, 6, 134-141.

Pollack, I., and Decker, L. R. Confidence ratings, message reception, and the receiver operating characteristic. J. Acoust. Soc. Amer., 1958, 30, 286-292.

Postman, L., and Rosenzweig, M. R. Practice and transfer in the visual and auditory recognition of verbal stimuli. Amer. J. Psychol., 1956, 69, 209-226.

Sarbin, T. R., Taft, R., and Bailey, D. E. Clinical inference and cognitive theory. New York: Holt-Rinehart, 1960.

Solomon, R. L., and Postman, L. Frequency of usage as a determinant of recognition thresholds for words. J. exp. Psychol., 1952, 43, 195-201.

Summers, S. A. Learning of responses to multiple weighted cues. J. exp. Psychol., 1962, 64, 29-34.

Swets, J. A. Is there a sensory threshold? Science, 1961, 134, 168-177.

Tanner, W. P., and Swets, J. A. A decision-making theory of visual detection. Psychol. Rev., 1954, 61, 401-409.

Todd, F. J. A methodological analysis of clinical judgment. Unpublished doctoral dissertation, Univ. of Colorado, 1954.

Todd, F. J. A methodological analysis of clinical judgment. Unpublished doctoral dissertation, Univ. of Colorado, 1954.

NOTES

1. A procedure for measurement of the distances between a $\underline{\emptyset}$ and the $\underline{\theta}$s in the space has not been formulated but a few suggestions are possible. For example, one might use a $\underline{D}^2$ (Cronbach and Glaser, 1953; Osgood and Suci, 1952) computed for the cue values and dimensions common to the $\underline{\emptyset}$ and each $\underline{\theta}$. Or, a technique of measuring subjective similarity recently developed by Ekman, Goude, and Waern (1961) might provide the necessary method.

2. The resemblance between this process and statistical decision theory is obvious. The decision paradigm, as well as the concept of the critical value fluctuating as a function of motivation, is based on a theory proposed by Tanner and Swets (1954) for sensory thresholds. (See also, Swets, 196L)

3. The question arises about the special case in which a number of $\underline{\theta}$s are all the same distance from the $\underline{\emptyset}$ and all lie within the critical range around it. $\underline{S}$s apparently resolve this situation by matching the relative frequency of their choices of each $\underline{\theta}$ with the relative frequency with which they have previously seen each $\underline{\theta}$ (Binder and Feldman, 1960). Of course this is an extremely artificial situation. Ordinarily, when faced with more than one correct $\underline{\theta}$, another cue value will be sought in order to break the deadlock. In the laboratory, however, the obtaining of additional cues is not permitted and $\underline{S}$ must fall back on another system, probability matching.

4. While familiarity is undoubtedly influenced by the duration of each encounter and the amount of elapsed time between encounters, we will consider frequency as its main determinant. Familiarity has been demonstrated to be a monotonic, negatively accelerated function of frequency of past encounter. (Arnoult, 1956; Nobel, 1954; and Solomon and Postman, 1952).

5. Note that, if $\underline{S}$ erroneously accepts that the $\underline{\emptyset}$ and a particular $\underline{\Theta}$ derive from the same $\underline{O}$, he will also err in attributing the $\underline{\Theta}$'s cue values to the $\underline{O}$. Like the trick cards in the experiments by Beach (1961, 1962) this source of error may provide a method for systematically investigating the influence of the distance between $\underline{\emptyset}$s and $\underline{\Theta}$s upon $\underline{S}$s' choice of inference method. For example, if $\underline{S}$ mistakenly attributes $\underline{\Theta_1}$'s set of unknown cue values to $\underline{O_2}$, it can be assumed that the distance between $\underline{\Theta_1}$ and $\underline{\emptyset_2}$ on the known cue dimensions is small enough to allow the $\underline{S}$ to use either recognition or assimilation rather th n identification.

6. This discussion concerns the case in which $\underline{S}$ makes one judgment for the value of the $\underline{\Theta}$'s attribute. If the situation is altered, as is done in probability learning experiments, and $\underline{S}$ is required to make a large number of corrected responses with the most recent value randomly varied on every trial, we would expect that the response distribution would generally resemble the objective frequency distribution for the various values of the attributes, except that the response distribution should be slightly lower than the objective frequency distribution for the less frequent values and slightly higher around the modal value, (Attneave, 1953; Gardner, 1957). For simplicity we have discussed only the mode and the most recent value of the attribute in determining the compromise. Research by Hake and Hyman (1953) indicates, however, that the previous <u>two</u> most recent events should probably be considered, with the oldest one weighted less. Any attempt to develop this notion further will have to take this fact into account.